Bernoulli **20**(3), 2014, 1620–1646
DOI: 10.3150/13-BEJ535

Fisher information and convergence to stable laws

S.G. BOBKOV[1], G.P. CHISTYAKOV[2] and F. GÖTZE[2]

[1]*School of Mathematics, University of Minnesota, 127 Vincent Hall, 206 Church St. S.E., Minneapolis, MN 55455, USA. E-mail: bobkov@math.umn.edu*
[2]*Faculty of Mathematics, University of Bielefeld, Postfach 100131, 33501 Bielefeld, Germany. E-mail:* [*]*chistyak@math.uni-bielefeld.de;* [**]*goetze@mathematik.uni-bielefeld.de*

The convergence to stable laws is studied in relative Fisher information for sums of i.i.d. random variables.

Keywords: Fisher information; limit theorems; stable laws

1. Introduction

Let $(X_n)_{n\geq 1}$ be independent identically distributed random variables. Define the normalized sums

$$Z_n = \frac{X_1 + \cdots + X_n}{b_n} - a_n$$

for given (non-random) normalizing sequences $a_n \in \mathbf{R}$ and $b_n > 0$. Assuming that Z_n converges weakly in distribution to a random variable Z with a non-degenerate stable law, we consider the Fisher information distance

$$I(Z_n \| Z) = \int_{-\infty}^{\infty} \left(\frac{p_n'(x)}{p_n(x)} - \frac{\psi'(x)}{\psi(x)} \right)^2 p_n(x)\,\mathrm{d}x,$$

where p_n and ψ denote the densities of Z_n and Z, respectively. The definition makes sense, if p_n is absolutely continuous and is supported on the support interval of ψ, with a Radon–Nikodym derivative $p_n'(x)$. Otherwise, put $I(Z_n \| Z) = \infty$.

If X_1 has finite second moment with mean zero and variance one, the classical central limit theorem is valid, that is, $Z_n \Rightarrow Z$ weakly in distribution, with $a_n = 0$, $b_n = \sqrt{n}$, where Z is standard normal. In this case a striking result of Barron and Johnson [8] indicates that $I(Z_n \| Z) \to 0$, as $n \to \infty$, as long as $I(Z_n \| Z) < \infty$, for some n, that is, if for some n, Z_n has finite Fisher information

$$I(Z_n) = \int_{-\infty}^{\infty} \frac{p_n'(x)^2}{p_n(x)}\,\mathrm{d}x.$$

This observation considerably strengthens a number of results on the central limit theorem for strong distances involving the total variation and the relative entropy. It raises at the same time the question about possible extensions to non-normal limit stable laws (as mentioned, e.g., in [7],

1350-7265 © 2014 ISI/BS

page 104). The question turns out to be rather tricky, and it is not that evident that $I(Z_n)$ needs to be even bounded for large n (a property which is guaranteed by Stam's inequality in case of a finite second moment).

The present note gives an affirmative solution of the problem in case of the so-called non-extremal stable laws, cf. Definition 1.2 below. In the sequel, we shall consider non-degenerate distributions, only.

Theorem 1.1. *Assume that the sequence of normalized sums Z_n defined above converges weakly in distribution to a random variable Z with a non-extremal stable limit law. Then $I(Z_n\|Z) \to 0$, as $n \to \infty$, if and only if $I(Z_n\|Z) < \infty$ for some n.*

The normal case is included in this assertion. Note, however, that if X_1 has an infinite second moment, but still belongs to the domain of normal attraction, we have $I(Z_n\|Z) = \infty$ for all n. Hence, in this special case there is no convergence in relative Fisher information.

In the remaining cases, Z has a stable distribution with some parameters $0 < \alpha < 2$, $-1 \le \beta \le 1$, with characteristic function $f(t) = \mathbf{E}e^{itZ}$ described by

$$f(t) = \exp\bigl\{iat - c|t|^\alpha\bigl(1 + i\beta\,\mathrm{sign}(t)\omega(t,\alpha)\bigr)\bigr\}, \tag{1.1}$$

where $a \in \mathbf{R}$, $c > 0$, and $\omega(t,\alpha) = \tan(\frac{\pi\alpha}{2})$ in case $\alpha \ne 1$, and $\omega(t,\alpha) = \frac{2}{\pi}\log|t|$ for $\alpha = 1$. In particular, $|f(t)| = e^{-c|t|^\alpha}$ which implies that Z has a smooth density $\psi(x)$.

Definition 1.2. *A stable distribution is called non-extremal, if it is normal or, if $0 < \alpha < 2$ and $-1 < \beta < 1$ in (1.1).*

In the latter case, the density ψ of Z is known to satisfy asymptotic relations

$$\psi(x) \sim c_0|x|^{-(1+\alpha)} \qquad (x \to -\infty), \qquad \psi(x) \sim c_1 x^{-(1+\alpha)} \qquad (x \to \infty) \tag{1.2}$$

with some constants $c_0, c_1 > 0$. Since any stable distribution is also unimodal (cf. [14]), ψ has to be positive on the whole real line, as follows from (1.2).

The property that X_1 belongs to the domain of attraction of a stable law of index $0 < \alpha < 2$ may be expressed explicitly in terms of the distribution function $F_1(x) = \mathbf{P}\{X_1 \le x\}$. Namely, we have $Z_n \Rightarrow Z$ with some $b_n > 0$ and $a_n \in \mathbf{R}$, if and only if

$$F_1(x) = \bigl(c_0 + o(1)\bigr)|x|^{-\alpha}B\bigl(|x|\bigr) \qquad (x \to -\infty), \tag{1.3}$$

$$1 - F_1(x) = \bigl(c_1 + o(1)\bigr)x^{-\alpha}B(x) \qquad (x \to \infty) \tag{1.4}$$

for some constants $c_0, c_1 \ge 0$ that are not both zero, and where $B(x)$ is a slowly varying function in the sense of Karamata. This description reflects a certain behaviour of the characteristic function $f_1(t) = \mathbf{E}e^{itX_1}$ near the origin (cf. [6,15]).

In connection with Theorem 1.1, let us note that a similar assertion has recently been proved in [4] for the relative entropy

$$D(Z_n\|Z) = \int_{-\infty}^{\infty} p_n(x)\log\frac{p_n(x)}{\psi(x)}\,dx,$$

called also the Kullback–Leibler distance form the distribution of Z_n to the distribution of Z. It is shown that $D(Z_n\|Z) \to 0$, if and only if $Z_n \Rightarrow Z$ and $D(Z_n\|Z) < \infty$ for some n. In the normal case this result is due to Barron [1], which in turn goes back to the work by Linnik [10], initiating an information-theoretic approach to the central limit theorem.

To compare with other strong types of convergence, in the normal case it is known that, if $\mathbf{E}X_1 = \mathbf{E}Z$ and $\mathrm{Var}(X_1) = \mathrm{Var}(Z) = \sigma^2$, then

$$\frac{\sigma^2}{2} I(Z_n\|Z) \geq D(Z_n\|Z) \geq \frac{1}{2}\|F_n - \Phi\|_{\mathrm{TV}}^2, \tag{1.5}$$

where $\|F_n - \Phi\|_{\mathrm{TV}}$ is the distance in total variation norm between the distributions of Z_n and Z (denoted here by F_n and Φ, resp.). The first relation in (1.5), due to Stam [13], may be viewed as an information theoretic variant of Gross' logarithmic Sobolev inequality for the Gaussian measure. The second one is a particular case of the Pinsker-type inequality in which normality of Z has no special role [5,9,11]. Hence, the convergence to the normal law in Fisher information distance is a stronger property than in total variation and even than in relative entropy. The question of how the Fisher information and entropic distances are related to each other with respect to other stable laws does not seem to have been addressed in the literature. Apparently it is a question about the existence of certain weak logarithmic Sobolev inequalities for probability distributions with heavy tails, and we do not touch it here. However, it is natural to conjecture that the situation is similar as in the normal case via a suitable analogue of (1.5).

Another obvious question concerns the description of distributions satisfying the conditions of Theorem 1.1. In the non-normal case, the property $I(Z_n\|Z) < \infty$ may be simplified to $I(Z_n) < \infty$. Taking, for example, $n = 1$, we obtain $I(X_1) < \infty$ as a sufficient condition, which is however rather strong and may be considerably weakened by choosing larger values of n. One may wonder therefore what assumptions need to be added to (1.3)–(1.4) in terms of F_1 or f_1 to obtain the convergence of Z_n to Z in relative Fisher information. As shown in [3], for some n, we have $I(Z_n) < \infty$, if and only if, for some n, Z_n has a continuously differentiable density p_n such that

$$\int_{-\infty}^{\infty} \left|p_n'(x)\right| \mathrm{d}x < \infty.$$

Still equivalently, for some n, p_n has to be a function of bounded variation. Moreover, if X_1 has a finite first absolute moment, this property may be formulated explicitly in terms of the behaviour of f_1 at infinity, as any of the following two equivalent assertions:

(a) For some $\varepsilon > 0$, $|f_1(t)| = \mathrm{O}(t^{-\varepsilon})$, as $t \to \infty$;

(b) For some $\nu > 0$,

$$\int_{-\infty}^{\infty} \left|f_1(t)\right|^{\nu} t^2 \, \mathrm{d}t < \infty. \tag{1.6}$$

This characterization may be used in Theorem 1.1 in case $1 < \alpha \leq 2$, since then, by (1.3)–(1.4), we have $\mathbf{E}|X_1|^{\delta} < \infty$, for all $0 < \delta < \alpha$.

Corollary 1.3. *Assume that the sequence Z_n as above converges weakly in distribution to a random variable Z with a non-extremal stable limit law with index $1 < \alpha \leq 2$. Then $I(Z_n\|Z) \to 0$, as $n \to \infty$, if and only if (1.6) holds for some $\nu > 0$.*

In particular, this description is applicable to the usual central limit theorem, that is, when X_1 has finite second moment. In this case (cf. [3]), (1.6) is equivalent to the formally weaker condition

$$\int_{-\infty}^{\infty} \left|f_1(t)\right|^{\nu} |t| \,\mathrm{d}t < \infty \qquad \text{for some } \nu > 0.$$

However, removing the weight $|t|$ from the above integral, we obtain an essentially weaker (so-called "smoothness") property

$$\int_{-\infty}^{\infty} \left|f_1(t)\right|^{\nu} \,\mathrm{d}t < \infty \qquad \text{for some } \nu > 0. \tag{1.7}$$

Once it is known that $Z_n \Rightarrow Z$ weakly in distribution with a stable limit law (for the i.i.d. summands as above), the condition (1.7) allows one to strengthen the weak convergence in the following sense. It is equivalent to the property that, for some and consequently for any sufficiently large n, Z_n has an absolutely continuous distribution with a bounded continuous density p_n. Moreover, in that and only that case, the uniform local limit theorem holds: $\sup_x |p_n(x) - \psi(x)| \to 0$, as $n \to \infty$ (cf. [6]).

The paper is organized as follows. First, we state some general bounds on Fisher information and some properties of densities which can be represented as convolutions of densities with finite Fisher information (Sections 2–4). A main result used here has been already proved in recent work [3]. In Section 5, we turn to the stable case and discuss a number of auxiliary results such as local limit theorems, as well as questions about the behaviour of characteristic functions of Z_n near zero. In Section 6, we reduce Theorem 1.1 to showing that the Fisher information $I(Z_n)$ is bounded in n. The subsequent sections are therefore focused on this boundedness problem. Section 7 introduces a special decomposition of convolutions, and the final steps of the proof of Theorem 1.1 can be found in Section 8. We shall complement the proofs by comments explaining why the condition (1.6) is sufficient for the validity of Theorem 1.1.

2. General results about Fisher information

Definition 2.1. *If a random variable X has an absolutely continuous density p with Radon–Nikodym derivative p', its Fisher information is defined by*

$$I(X) = I(p) = \int_{\{p(x)>0\}} \frac{p'(x)^2}{p(x)} \,\mathrm{d}x. \tag{2.1}$$

In this case, if $\tilde{p}(x) = p(x)$ for almost all x (a.e.), put $I(\tilde{p}) = I(p)$. In any other case, $I(X) = \infty$.

The equality (2.1) appears as a particular case of the Fisher information

$$J(\theta) = \int_{-\infty}^{\infty} \left(\frac{\partial p_\theta(x)}{\partial \theta}\right)^2 p_\theta(x) \,\mathrm{d}x$$

for the family of densities $p_\theta(x) = p(x-\theta)$ with respect to the location parameter $\theta \in \mathbf{R}$.

If $I(X)$ as defined in (2.1) is finite, then necessarily the distribution of X has to be absolutely continuous with density $p(x)$ such that the derivative $p'(x)$ exists and is finite on a set of full Lebesgue measure (and then p will always be chosen to be a.e. differentiable). Furthermore, one can show that, if $I(X) < \infty$, then $p'(x) = 0$ at any point, where $p(x) = 0$ (cf. [3]). With this in mind, the integration in (2.1) may be extended to the whole real line.

It follows immediately from the definition that the I-functional is translation invariant and homogeneous of order -2, that is, $I(a+bX) = \frac{1}{b^2} I(X)$, for all $a \in \mathbf{R}$ and $b \neq 0$.

Since the function u^2/v is convex in the upper half-plane $u \in \mathbf{R}$, $v > 0$, this functional is convex. That is, for all densities $p_1, \dots, p_n$, we have Jensen's inequality

$$I(\alpha_1 p_1 + \cdots + \alpha_n p_n) \le \sum_{k=1}^n \alpha_k I(p_k) \qquad \Biggl(\alpha_k > 0, \sum_{k=1}^n \alpha_k = 1\Biggr).$$

The inequality may be generalized to arbitrary "continuous" mixtures of densities. In particular, for the convolution

$$p * q(x) = \int_{-\infty}^{\infty} p(x-y) q(y) \, \mathrm{d}x$$

of any two densities p and q, we have

$$I(p * q) \le \min\bigl\{I(p), I(q)\bigr\}. \tag{2.2}$$

In other words, if X and Y are independent random variables with these densities, then

$$I(X+Y) \le \min\bigl\{I(X), I(Y)\bigr\}.$$

This property may be viewed as monotonicity of the Fisher information: this functional decreases when adding an independent summand. In fact, a much stronger inequality is available.

Proposition 2.2 (Stam [13]). *If X and Y are independent random variables, then*

$$\frac{1}{I(X+Y)} \ge \frac{1}{I(X)} + \frac{1}{I(Y)}. \tag{2.3}$$

Let us also introduce the Fisher information distance

$$I(X \| Z) = \int_{-\infty}^{+\infty} \left(\frac{p'(x)}{p(x)} - \frac{\psi'(x)}{\psi(x)}\right)^2 p(x) \, \mathrm{d}x$$

with respect to a random variable Z having a stable law. We need the following elementary observation, which shows that the question of boundedness of the Fisher information $I(Z_n)$ and of the Fisher information distance $I(Z_n \| Z)$ for the normalized sums Z_n as introduced in Theorem 1.1 are in fact equivalent.

Proposition 2.3. *If Z has a non-extremal stable law of some index $0 < \alpha < 2$, then for any random variable X,*

$$I(X\|Z) \le 2I(X) + c(Z), \tag{2.4}$$

$$I(X) \le 2I(X\|Z) + c(Z), \tag{2.5}$$

where $c(Z)$ depends on the distribution of Z, only. In particular, $I(X\|Z) < \infty$, if and only if $I(X) < \infty$.

Proof. The assertion is based on the fact that any non-extremal non-normal stable distribution has a smooth positive density ψ such that, for all $k = 1, 2, \dots,$

$$\bigl|\bigl(\log\psi(x)\bigr)^{(k)}\bigr| \sim \frac{(k-1)!}{|x|^k} \qquad \bigl(|x| \to \infty\bigr)$$

(cf. [6,15]). In particular, $\frac{|\psi'(x)|}{\psi(x)} \sim \frac{1}{|x|}$, so

$$\frac{|\psi'(x)|}{\psi(x)} \le \frac{c}{1+|x|} \qquad (x \in \mathbf{R}) \tag{2.6}$$

with some positive constant c (and the converse inequality is also true with positive constant for all large $|x|$). Hence, assuming that $I(X) < \infty$, then writing

$$\left(\frac{p'(x)}{p(x)} - \frac{\psi'(x)}{\psi(x)}\right)^2 \le 2\left(\frac{p'(x)}{p(x)}\right)^2 + 2\left(\frac{\psi'(x)}{\psi(x)}\right)^2 \le 2\left(\frac{p'(x)}{p(x)}\right)^2 + 2c^2$$

and integrating this inequality with weight $p(x)$, we obtain (2.4). Similarly,

$$\left(\frac{p'(x)}{p(x)}\right)^2 \le 2\left(\frac{p'(x)}{p(x)} - \frac{\psi'(x)}{\psi(x)}\right)^2 + 2c^2,$$

which leads to (2.5). □

Similar arguments for the normal case ($\alpha = 2$) however lead to a different conclusion. Indeed, if $Z \sim N(a, \sigma^2)$, we have $\frac{\psi'(x)}{\psi(x)} = -\frac{x-a}{\sigma^2}$, and we get the following proposition.

Proposition 2.4. *If Z is normal, then $I(X\|Z) < \infty$, if and only if $I(X) < \infty$ and $\mathbf{E}X^2 < \infty$.*

Note that in case where X and Z have equal means and variances, we have $I(X\|Z) = I(X) - I(Z)$.

3. Connection with functions of bounded variation

Applying Cauchy's inequality and using the remark that $p(x) = 0 \Rightarrow p'(x) = 0$ a.e., one immediately obtains from Definition 2.1 the following elementary lower bound on the Fisher information.

Proposition 3.1. *If X has an absolutely continuous density p with Radon–Nikodym derivative p', then*

$$\int_{-\infty}^{\infty} \left|p'(x)\right| \mathrm{d}x \le \sqrt{I(X)}. \tag{3.1}$$

Here, the integral represents the total variation norm of the function p as used in Real Analysis,

$$\|p\|_{\mathrm{TV}} = \sup \sum_{k=1}^{n} \left|p(x_k) - p(x_{k-1})\right|,$$

where the supremum runs over all finite collections $x_0 < x_1 < \cdots < x_n$.

The densities p with finite total variation are vanishing at infinity and are uniformly bounded by $\|p\|_{\mathrm{TV}}$. Moreover, their characteristic functions

$$f(t) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}tx} p(x)\,\mathrm{d}x \qquad (t \in \mathbf{R})$$

admit, by integration by parts, a simple upper bound

$$\left|f(t)\right| \le \frac{\|p\|_{\mathrm{TV}}}{|t|} \qquad (t \ne 0). \tag{3.2}$$

Hence, by Proposition 3.1, if a random variable X has finite Fisher information, its density p and characteristic function $f(t) = \mathbf{E}\mathrm{e}^{\mathrm{i}tX}$ satisfy similar bounds

$$\sup_x p(x) \le \sqrt{I(X)}, \qquad \left|f(t)\right| \le \frac{\sqrt{I(X)}}{|t|} \qquad (t \ne 0). \tag{3.3}$$

In general, the inequality (3.1) cannot be reversed, though this is possible for convolutions of three densities of bounded variation. The following statement may be found in [3].

Proposition 3.2. *If independent random variables X_j $(j = 1, 2, 3)$ have densities p_j of bounded variation, then $S = X_1 + X_2 + X_3$ has finite Fisher information, and moreover,*

$$I(S) \le \tfrac{1}{2}\left[\|p_1\|_{\mathrm{TV}}\|p_2\|_{\mathrm{TV}} + \|p_1\|_{\mathrm{TV}}\|p_3\|_{\mathrm{TV}} + \|p_2\|_{\mathrm{TV}}\|p_3\|_{\mathrm{TV}}\right]. \tag{3.4}$$

Note that the convolution of two densities of bounded variation may have an infinite Fisher information. For example, the convolution of the uniform distribution on $(-\frac{1}{2}, \frac{1}{2})$ with itself has the triangle density $p(x) = \max(1 - |x|, 0)$, in which case $I(p) = \infty$.

Remark 3.3. A similar bound on the Fisher information may also be given in terms of characteristic functions. In view of (3.4), it suffices to bound the total variation norm, and this can be done by applying the inverse Fourier formula, at least in case of finite first absolute moment.

One can show that, if the characteristic function $f(t)$ of a random variable X is continuously differentiable for $t > 0$, and

$$\int_{-\infty}^{\infty} t^2\big(|f(t)|^2 + |f'(t)|^2\big)\,\mathrm{d}t < \infty, \tag{3.5}$$

then X must have an absolutely continuous distribution with density p of bounded total variation satisfying

$$\|p\|_{\mathrm{TV}} \le \left(\int_{-\infty}^{\infty} |tf(t)|^2\,\mathrm{d}t \int_{-\infty}^{\infty} \big|\big(tf(t)\big)'\big|^2\,\mathrm{d}t\right)^{1/4}. \tag{3.6}$$

We refer to [3] for details.

4. Classes of densities representable as convolutions

General bounds like (3.3) may considerably be sharpened in the case where p is representable as convolution of several densities with finite Fisher information. Here, we consider the collection $\mathfrak{P}_2(I)$ of all functions on the real line which can be represented as convolution of two probability densities with Fisher information at most I. Correspondingly, let $\mathfrak{P}_2 = \bigcup_I \mathfrak{P}_2(I)$ denote the collection of all functions representable as convolution of two probability densities with finite Fisher information. Note that, by (2.3), $I(p) \le \frac{1}{2}I$, for any $p \in \mathfrak{P}_2(I)$.

Thus, a random variable $X = X_1 + X_2$ has density p in $\mathfrak{P}_2$, if it may be written as

$$p(x) = \int_{-\infty}^{\infty} p_1(x-y)p_2(y)\,\mathrm{d}x \tag{4.1}$$

in terms of absolutely continuous densities p_1, p_2 of the independent summands X_1, X_2 having finite Fisher information. Differentiating under the integral sign, we obtain a Radon–Nikodym derivative of the function p,

$$p'(x) = \int_{-\infty}^{\infty} p_1'(x-y)p_2(y)\,\mathrm{d}y = \int_{-\infty}^{\infty} p_1'(y)p_2(x-y)\,\mathrm{d}y. \tag{4.2}$$

The latter expression shows that p' is an absolutely continuous function and has the Radon–Nikodym derivative

$$p''(x) = \int_{-\infty}^{\infty} p_1'(y)p_2'(x-y)\,\mathrm{d}y. \tag{4.3}$$

In other words, p'' appears as the convolution of the functions p_1' and p_2' which are integrable, according to Proposition 3.1.

Note that equality (4.3) defines $p''(x)$ at every individual point x, not just almost everywhere (which is typical for a Radon–Nikodym derivative). Using the property $p_j(x) = 0 \Rightarrow p_j'(x) = 0$ in case of finite Fisher information, we obtain a similar implication $p(x) = 0 \Rightarrow p''(x) = 0$, which holds for any x.

Moreover, since by (4.3),

$$\left|p''(x)\right| \le \int_{-\infty}^{\infty} \left|p_1'(y)\right|\left|p_2'(x-y)\right| \mathrm{d}y,$$

a direct application of the inequality (3.1) together with Fubini's theorem shows that p' has finite total variation

$$\|p'\|_{\mathrm{TV}} = \int_{-\infty}^{\infty} \left|p''(x)\right| \mathrm{d}x \le I.$$

These formulas may be used to derive various pointwise and integral relations within the class $\mathfrak{P}_2$ such as the following statement (which also summarizes the previous remarks).

Proposition 4.1. *Any density p in $\mathfrak{P}_2(I)$ has an absolutely continuous derivative p' of bounded variation satisfying, for all $x \in \mathbf{R}$,*

$$\left|p'(x)\right| \le I^{3/4}\sqrt{p(x)} \le I. \tag{4.4}$$

In addition,

$$\int_{-\infty}^{\infty} \frac{p''(x)^2}{p(x)}\, \mathrm{d}x \le I^2. \tag{4.5}$$

To be more precise, integration in (4.5) is restricted to the set $\{p(x) > 0\}$. This proposition can be found in [3]; since the proof is short, we shall include it here for completeness.

Proof of Proposition 4.1. Starting with the representations (4.1)–(4.2), in which $I(p_j) \le I$, define the functions $u_j(x) = \frac{p_j'(x)}{\sqrt{p_j(x)}} 1_{\{p_j(x)>0\}}$ $(j = 1, 2)$. Applying Cauchy's inequality, we get

$$\begin{aligned}
p'(x)^2 &= \left(\int_{-\infty}^{\infty} u_1(x-y) \cdot \sqrt{p_1(x-y)}\, p_2(y)\, \mathrm{d}y\right)^2 \\
&\le I(X_1) \int_{-\infty}^{\infty} p_1(x-y) p_2(y)^2\, \mathrm{d}y \\
&\le I(X_1)\sqrt{I(X_2)} \int_{-\infty}^{\infty} p_1(x-y) p_2(y)\, \mathrm{d}y = I(X_1)\sqrt{I(X_2)}\, p(x),
\end{aligned}$$

where we used $p_2(y) \le \sqrt{I(X_2)}$, according to (3.3). Hence, we obtain the first inequality in (4.4), and the second follows from $p(x) \le \sqrt{I}$. Similarly, rewrite (4.3) as

$$p''(x) = \int_{-\infty}^{\infty} \big(u_1(x-y) u_2(y)\big) \sqrt{p_1(x-y) p_2(y)}\, \mathrm{d}y$$

to get

$$p''(x)^2 \le \int_{-\infty}^{\infty} u_1(x-y)^2 u_2(y)^2\, \mathrm{d}y \int_{-\infty}^{\infty} p_1(x-y) p_2(y)\, \mathrm{d}y = u(x)^2 p(x),$$

where we define $u \geq 0$ by

$$u(x)^2 = \int_{-\infty}^{\infty} u_1(x-y)^2 u_2(y)^2 \, \mathrm{d}y.$$

It follows that

$$\int_{-\infty}^{\infty} u(x)^2 \, \mathrm{d}x = I(X_1)I(X_2) \leq I^2,$$

which implies (4.5). □

The analytic properties of densities in $\mathfrak{P}_2$ allow us to make use of different formulas for the Fisher information (by using integration by parts). For example,

$$I(X) = -\int_{-\infty}^{\infty} p''(x) \log p(x) \, \mathrm{d}x,$$

provided that the integrand is Lebesgue integrable.

We will need the following "tail-type" estimate for the Fisher information.

Corollary 4.2. *If p is in $\mathfrak{P}_2(I)$, then for any T real,*

$$\int_T^{\infty} \frac{p'(x)^2}{p(x)} \, \mathrm{d}x \leq I^{3/4} \sqrt{p(T)} \bigl|\log p(T)\bigr| + I \left(\int_T^{\infty} p(x) \log^2 p(x) \, \mathrm{d}x \right)^{1/2}. \tag{4.6}$$

Proof. Assuming that the last integral is finite, let us decompose the open set $G = \{x > T: p(x) > 0\}$ into the union of at most countably many disjoint intervals (a_n, b_n), $T \leq a_n < b_n \leq \infty$.

If $a_n > T$, we have $p(a_n) = 0$, so $p'(x) \log p(x) \to 0$, as $x \downarrow a_n$, by Proposition 4.1. Similarly, $p(b_n) = 0$, if $b_n < \infty$, and in addition $p(\infty) = 0$.

Let $a_n < T_1 < T_2 < b_n$. Since p' is an absolutely continuous function of bounded variation, integration by parts is justified and yields

$$\int_{T_1}^{T_2} \frac{p'(x)^2}{p(x)} \, \mathrm{d}x = \int_{T_1}^{T_2} p'(x) \, \mathrm{d} \log p(x) = p'(x) \log p(x) \Big|_{x=T_1}^{T_2} - \int_{T_1}^{T_2} p''(x) \log p(x) \, \mathrm{d}x.$$

Letting $T_1 \to a_n$ and $T_2 \to b_n$, we get in case $a_n > T$

$$\int_{a_n}^{b_n} \frac{p'(x)^2}{p(x)} \, \mathrm{d}x = -\int_{a_n}^{b_n} p''(x) \log p(x) \, \mathrm{d}x$$

and

$$\int_{a_n}^{b_n} \frac{p'(x)^2}{p(x)} \, \mathrm{d}x = -p'(T) \log p(T) - \int_{a_n}^{b_n} p''(x) \log p(x) \, \mathrm{d}x$$

in case $a_n = T$ (if such n exists). Anyhow, the summation over n gives

$$\int_G \frac{p'(x)^2}{p(x)}\,dx \le \left|p'(T)\log p(T)\right| + \int_G \left|p''(x)\log p(x)\right|dx. \tag{4.7}$$

Here the first term on the right-hand side can be estimated by virtue of (4.4), which leads to the first term on the right-hand side of (4.6). Using (4.5) together with Cauchy's inequality, for the last integral we also have

$$\left(\int_G \frac{|p''(x)|}{\sqrt{p(x)}}\sqrt{p(x)}\left|\log p(x)\right|dx\right)^2 \le I^2 \int_T^\infty p(x)\log^2 p(x)\,dx,$$

thus proving Corollary 4.2. □

5. Stable laws and uniform local limit theorems

Let us return to the normalized sums

$$Z_n = \frac{1}{b_n}(X_1 + \cdots + X_n) - a_n \qquad (a_n \in \mathbf{R}, b_n > 0),$$

associated with independent identically distributed random variables $(X_n)_{n\ge 1}$. In this section, we discuss uniform limit theorems for densities p_n of Z_n and behaviour of their characteristic functions near the origin. As before, if $Z_n \Rightarrow Z$, the density and the characteristic function of the stable limit Z are denoted by ψ and f, respectively.

Introduce the characteristic functions of X_1 and Z_n,

$$f_1(t) = \mathbf{E}e^{itX_1}, \qquad f_n(t) = \mathbf{E}e^{itZ_n} = e^{-ita_n} f_1(t/b_n)^n \qquad (t \in \mathbf{R}).$$

To avoid confusion, we make the convention that $Z_1 = X_1$, that is, $a_1 = 0$ and $b_1 = 1$.

Proposition 5.1. *Assume that $Z_n \Rightarrow Z$ weakly in distribution. If*

$$\int_{-\infty}^{\infty} \left|f_1(t)\right|^\nu dt < \infty \qquad \textit{for some } \nu > 0, \tag{5.1}$$

then for all n large enough, Z_n have bounded continuous densities p_n such that

$$\lim_{n\to\infty} \sup_x \left|p_n(x) - \psi(x)\right| = 0. \tag{5.2}$$

Proposition 5.2. *Assume that $Z_n \Rightarrow Z$ weakly in distribution. If*

$$\int_{-\infty}^{\infty} \left|f_1(t)\right|^\nu |t|\,dt < \infty \qquad \textit{for some } \nu > 0, \tag{5.3}$$

then for all n large enough, Z_n have continuously differentiable densities p_n with bounded derivatives, and moreover

$$\lim_{n\to\infty} \sup_x \bigl|p_n'(x) - \psi'(x)\bigr| = 0. \tag{5.4}$$

The first assertion is well known, cf. [6], page 126. The condition (5.1) is actually equivalent to the property that for all sufficiently large n, say $n \ge n_0$, Z_n have bounded continuous densities p_n. In that case, the characteristic functions f_n are integrable whenever $n \ge 2n_0$. Conversely, under (5.1), these densities for $n \ge \nu$ are given by the inversion formula

$$p_n(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}tx} f_n(t)\,\mathrm{d}t. \tag{5.5}$$

Under the stronger assumption (5.3), the above equality may be differentiated, and we get a similar representation for the derivative

$$p_n'(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (-\mathrm{i}t)\mathrm{e}^{-\mathrm{i}tx} f_n(t)\,\mathrm{d}t. \tag{5.6}$$

Although Proposition 5.2 is not stated in [6], its proof is similar to the proof of Proposition 5.1. An important ingredient in the argument is the fact that the weak convergence $Z_n \Rightarrow Z$ forces f_1 to be regularly behaving near the origin. This fact can also be used in the study of the boundedness of the Fisher information distance $I(Z_n \| Z)$, so let us state it separately.

Proposition 5.3. *Let $Z_n \Rightarrow Z$ weakly in distribution, where Z has a stable law of index $0 < \alpha < 2$. Then*

$$\bigl|f_1(t)\bigr| = \exp\bigl\{-c|t|^{\alpha} h\bigl(1/|t|\bigr)\bigr\}, \tag{5.7}$$

where $c > 0$ and $h(x)$ is a slowly varying function for $x \to \infty$ such that

$$\lim_{n\to\infty} \frac{nh(b_n)}{b_n^{\alpha}} = 1. \tag{5.8}$$

Moreover, there is a constant $c > 0$ such that, as $n \to \infty$,

$$\mathbf{P}\bigl\{|X_1| > b_n\bigr\} \sim \frac{c}{n}. \tag{5.9}$$

In comparison with (5.7) a more precise statement is obtained in [6], cf. Theorem 2.6.5, page 85. Namely, if $Z_n \Rightarrow Z$, where Z has a stable distribution of index $0 < \alpha < 2$, then for all t small enough,

$$f_1(t) = \exp\bigl\{\mathrm{i}\gamma t - c|t|^{\alpha} h\bigl(1/|t|\bigr)\bigl(1 + \mathrm{i}\beta\, \mathrm{sign}(t)\omega(t,\alpha)\bigr)\bigr\},$$

where γ is real, $c > 0$, and the parameter $\beta \in [-1, 1]$ and the function $\omega(t, \alpha)$ are the same as in the representation (1.1) for the characteristic function $f(t)$ of Z. By lengthy computations in the proof of Theorem 2.6.5 in [6], it was shown that the function $B(x)$ appearing in the asymptotic relations (1.3)–(1.4) and the function $h(x)$ are connected via

$$h(x) = \bigl(1 + \mathrm{o}(1)\bigr) B(x) \qquad \text{as } x \to \infty.$$

Taking into account (5.8), this yields (5.9).

Remark. As shown in [6], the representation (5.7) together with the relation (5.8) remain to hold for $\alpha = 2$, that is, when Z is normal. Note that, if $\mathbf{E}X_1^2 < \infty$, one may take $h(x) = 1$ and $b_n \sim \sqrt{n}$. In that case, $\mathbf{P}\{|X_1| > b_n\} = \mathrm{o}(\frac{1}{n})$, as $n \to \infty$, so (5.9) is no longer true.

Let us return to the local limit theorems.

Proof of Proposition 5.2. From (5.6), we obtain the representation

$$p_n'(x) - \psi'(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} (-\mathrm{i}t) \mathrm{e}^{-\mathrm{i}tx} \bigl(f_n(t) - f(t)\bigr) \, \mathrm{d}t.$$

As is standard, we split the last integral into the three parts L_1, L_2, L_3 corresponding to integration over the regions $|t| \le T_n$, $T_n < |t| < T_n'$ and $|t| \ge T_n'$, respectively.

By the weak convergence, $f_n(t) \to f(t)$ uniformly on all intervals, and moreover,

$$\delta_n = \max_{|t| \le T_n} \bigl| f_n(t) - f(t) \bigr| \to 0 \qquad \text{as } n \to \infty$$

for some $T_n \to \infty$. Hence,

$$|L_1| = \left| \int_{|t| \le T_n} (-\mathrm{i}t) \mathrm{e}^{-\mathrm{i}tx} \bigl(f_n(t) - f(t)\bigr) \, \mathrm{d}t \right| \le \delta_n T_n^2 \to 0,$$

provided that T_n grows to infinity sufficiently slowly (which may be assumed).

Now, one of the consequences of (5.7), using the above remark about the normal case, is that, given $0 < \delta < \alpha$, the characteristic functions f_n admit on a relatively large interval the bound

$$\bigl| f_n(t) \bigr| \le \mathrm{e}^{-c(\delta)|t|^\delta} \qquad \bigl(|t| \le \varepsilon b_n\bigr) \tag{5.10}$$

with some positive constants ε and $c(\delta)$ which are independent of n, cf. [6], page 123. A similar bound holds for $f(t)$ itself, which is also seen from the representation (1.1). Hence, choosing $T_n' = \varepsilon b_n$, we have

$$\begin{aligned} |L_2| &= \left| \int_{T_n < |t| < T_n'} (-\mathrm{i}t) \mathrm{e}^{-\mathrm{i}tx} \bigl(f_n(t) - f(t)\bigr) \, \mathrm{d}t \right| \\ &\le 2 \int_{|t| > T_n} |t| \mathrm{e}^{-c(\delta)|t|^\delta} \, \mathrm{d}t \to 0. \end{aligned}$$

Finally, put $c = \sup_{|t|\geq\varepsilon} |f_1(t)|$. The condition (5.3) ensures that $f_1(t) \to 0$, as $t \to \infty$, so $c < 1$. Hence, for all $n \geq \nu$,

$$\int_{|t|\geq T_n'} |t| \left|f_n(t)\right| \mathrm{d}t = b_n^2 \int_{|t|\geq\varepsilon} |t| \left|f_1(t)\right|^n \mathrm{d}t$$

$$\leq b_n^2 c^{n-\nu} \int_{|t|\geq\varepsilon} |t| \left|f_1(t)\right|^\nu \mathrm{d}t \to 0.$$

Thus, $L_3 \to 0$, as well. □

From (5.2) and (5.4), we immediately obtain the convergence of a "truncated" Fisher information distance.

Corollary 5.4. *Assume that $Z_n \Rightarrow Z$ weakly in distribution, where Z has a non-extremal stable law. If $I(Z_{n_0}) < \infty$ for some n_0, then for all n large enough, the random variables Z_n admit continuously differentiable densities p_n, and for every fixed $T > 0$,*

$$\int_{-T}^{T} \left(\frac{p_n'(x)}{p_n(x)} - \frac{\psi'(x)}{\psi(x)}\right)^2 p_n(x)\,\mathrm{d}x = \mathrm{o}(1), \qquad n \to \infty. \tag{5.11}$$

Recall that the densities ψ of non-extremal stable laws are everywhere positive, which is the only additional property needed to show (5.11) on the basis of (5.2) and (5.4).

Indeed, by the assumption, we have $I(Z_n) < \infty$, for all $n \geq n_0$, and by (3.3),

$$\left|f_{n_0}(t)\right| \leq \frac{c}{|t|} \qquad (t \neq 0)$$

with $c = \sqrt{I(Z_{n_0})}$. Hence, the condition (5.3) is fulfilled with $\nu = 3n_0$. Therefore, we get both (5.2) and (5.4), and in particular, $p_n(x) \geq \varepsilon > 0$ in $|x| \leq T$, for all n large enough. As a result, the integrand in (5.11) is uniformly bounded from above by a sequence tending to zero.

6. Moderate deviations

As before, for independent identically distributed random variables $(X_n)_{n\geq 1}$, put

$$Z_n = \frac{X_1 + \cdots + X_n}{b_n} - a_n \qquad (a_n \in \mathbf{R}, b_n > 0). \tag{6.1}$$

It is well known that if $Z_n \Rightarrow Z$, where Z has a stable law of some index $0 < \alpha \leq 2$, then necessarily

$$b_n = n^{1/\alpha} h(n), \tag{6.2}$$

where h is a slowly varying function in the sense of Karamata.

To study the behaviour of $I(Z_n\|Z)$ in the non-extremal non-normal case, it is worthwhile noting that this Fisher information distance is finite, if and only if $I(Z_n)$ is finite (Proposition 2.3). In the normal case, $I(Z_n\|Z)<\infty$, if and only if $I(Z_n)<\infty$ and $\mathbf{E}Z_n^2<\infty$ (Proposition 2.4). The latter is equivalent to $\mathbf{E}X_1^2<\infty$, and then for the weak convergence $Z_n\Rightarrow Z$ with a standard normal limit one may take $b_n=\sqrt{n\operatorname{Var}X_1}$ and $a_n=\mathbf{E}X_1\sqrt{n}/\sqrt{\operatorname{Var}X_1}$.

In any case, the requirement that $I(Z_{n_0})<\infty$ implies that for all $n\geq n_0$, Z_n have absolutely continuous bounded densities which we denote in the sequel by p_n. Moreover, $p_n\in\mathfrak{P}_2$ whenever $n\geq 2n_0$, and then, by Proposition 4.1, p_n have continuous derivatives p_n' of bounded variation.

As the next step towards Theorem 1.1, we prove the following lemma.

Lemma 6.1. *Assume that $Z_n\Rightarrow Z$ weakly in distribution, where Z has a non-extremal stable law. If $\limsup_{n\to\infty}I(Z_n)<\infty$, then*

$$\lim_{n\to\infty}I(Z_n\|Z)=0. \tag{6.3}$$

Proof. As before, denote by ψ the density of Z, and put $S_n=X_1+\cdots+X_n$.

By the assumptions, $I'=\sup_{n\geq n_0}I(Z_n)<\infty$ for some n_0, so

$$I(S_n)\leq I'b_n^2,\qquad n\geq n_0.$$

If $n\geq 2n_0$, write $n=n_1+n_2$ with $n_1=[\frac{n}{2}]$, $n_2=n-n_1$. Then $n_1\geq n_0$ and $n_2\geq n_0$, and hence

$$I(S_{n_1})\leq I'b_{n_1}^2\leq Ib_n^2,\qquad I(S_n-S_{n_1})\leq I'b_{n_2}^2\leq Ib_n^2$$

with some constant I in view of the almost polynomial behaviour of b_n as described in (6.2). Thus,

$$Z_n=\left(\frac{S_{n_1}}{b_n}-a_n\right)+\frac{S_n-S_{n_1}}{b_n}$$

represents the sum of two independent random variables with Fisher information at most I. Therefore, $p_n\in\mathfrak{P}_2(I)$, for all $n\geq 2n_0$, and we may invoke Corollary 4.2.

In view of Corollary 5.4 we only need to show that, given $\varepsilon>0$, one may choose $T>0$ such that the integral

$$J=\int_{|x|>T}\left(\frac{p_n'(x)}{p_n(x)}-\frac{\psi'(x)}{\psi(x)}\right)^2p_n(x)\,\mathrm{d}x$$

is smaller than ε, for all n large enough.

Clearly, $J\leq 2J_1+2J_2$, where

$$J_1=\int_{|x|>T}\frac{p_n'(x)^2}{p_n(x)}\,\mathrm{d}x,\qquad J_2=\int_{|x|>T}\left(\frac{\psi'(x)}{\psi(x)}\right)^2p_n(x)\,\mathrm{d}x.$$

Recall that in case $0 < \alpha < 2$, we have $\frac{|\psi'(x)|}{\psi(x)} \leq \frac{c}{1+|x|}$ with a constant c depending on ψ, only (cf. (2.6)). Hence,

$$J_2 \leq \left(\frac{c}{1+T}\right)^2,$$

which thus can be made as small, as we wish.

If $\alpha = 2$ and $\mathbf{E}X_1^2 < \infty$, assume without loss of generality that $\mathbf{E}X_1 = 0$, $\mathbf{E}X_1^2 = 1$, so that ψ is a standard normal density, and

$$J_2 = \int_{|x|>T} x^2 p_n(x)\,\mathrm{d}x.$$

To bound these integrals, we appeal to the well-known large deviation relation

$$\mathbf{P}\big\{|\xi| \geq T\big\} \leq T \int_0^{2/T} \big(1 - \operatorname{Re} f(t)\big)\,\mathrm{d}t,$$

holding true for any random variable ξ with characteristic function $f(t)$. If $\mathbf{E}\xi^2 = 1$, and F is the distribution function of ξ, one may apply the same bound to the probability measure $x^2\,\mathrm{d}F(x)$ on the real line, and then it yields

$$\int_{|x|\geq T} x^2\,\mathrm{d}F(x) \leq T \int_0^{2/T} \big(1 + \operatorname{Re} f''(t)\big)\,\mathrm{d}t.$$

Hence,

$$J_2 \leq T \int_0^{2/T} \big(1 + \operatorname{Re} f_n''(t)\big)\,\mathrm{d}t,$$

where f_n denote the characteristic functions of Z_n. But, letting $g(t) = \mathrm{e}^{-t^2/2}$, as a variant of the central limit theorem, for any $c > 0$, one has $\sup_{|t|\leq c} |f_n''(t) - g''(t)| \to 0$, as $n \to \infty$, while $1 + g''(t) \to 0$, as $t \to 0$. This shows that, for T and n large enough, J_2 will be smaller than any prescribed positive number.

It remains to estimate J_1. We now apply (4.6) giving

$$\begin{aligned} J_1 &\leq I^{3/4}\big(\sqrt{p_n(T)}\big|\log p_n(T)\big| + \sqrt{p_n(-T)}\big|\log p_n(-T)\big|\big) \\ &\quad + 2I\bigg(\int_{|x|\geq T} p_n(x) \log^2 p_n(x)\,\mathrm{d}x\bigg)^{1/2}. \end{aligned} \tag{6.4}$$

Using the uniform local limit theorem in the form (5.2) together with the asymptotic relation (1.2) for $\psi(x)$ at infinity, we easily get

$$\sqrt{p_n(\pm T)}\big|\log p_n(\pm T)\big| \leq c\frac{\log T}{\sqrt{T}} + \varepsilon_n, \tag{6.5}$$

which holds for all sufficiently large n and all $T \ge T_0$ with $\varepsilon_n \to 0$ (as $n \to \infty$) and with constants $c > 0$ and T_0 large enough (depending on ψ, only).

To bound the integral in (6.4), we partition $\{x\colon |x| \ge T\}$ into the set

$$A = \{x\colon |x| \ge T,\, p_n(x) \le |x|^{-4}\}$$

and its complement B. By the definition,

$$\int_A p_n(x) \log^2 p_n(x)\,\mathrm{d}x \le 16 \int_{|x|\ge T} |x|^{-4} \log^2 |x|\,\mathrm{d}x \le \frac{32}{T}. \tag{6.6}$$

On the other hand, p_n are uniformly bounded, namely, $\sup p_n(x) \le \sqrt{I}$, for all $n \ge 2n_0$ (cf. (3.3)). Hence, on the set B,

$$\left|\log p_n(x)\right| \le \log \frac{\sqrt{I}}{p_n(x)} + |\log \sqrt{I}| \le 4 \log |x| + |\log I|$$

and therefore

$$\int_B p_n(x) \log^2 p_n(x)\,\mathrm{d}x \le c \int_{|x|\ge T} p_n(x) \log^2 |x|\,\mathrm{d}x, \tag{6.7}$$

where the constant depends on I.

Finally, we use the property that the moments $\mathbf{E}|Z_n|^\delta$ are uniformly bounded in n, whenever $0 < \delta < \alpha$ (cf. [6], page 142). Choosing $\delta = \alpha/2$ and using an elementary bound $|x|^{\alpha/4} \ge c_\alpha \log^2 |x|$ for $|x| \ge T_0$, we obtain with some constant K that

$$\begin{aligned} K &\ge \mathbf{E}|Z_n|^{\alpha/2} \ge T^{\alpha/4}\mathbf{E}|Z_n|^{\alpha/4} 1_{\{|Z_n|\ge T\}} \\ &= T^{\alpha/4} \int_{|x|\ge T} |x|^{\alpha/4} p_n(x)\,\mathrm{d}x \\ &\ge c_\alpha T^{\alpha/4} \int_{|x|\ge T} p_n(x) \log^2 |x|\,\mathrm{d}x. \end{aligned}$$

Thus, the second integral in (6.7) may be bounded by $cT^{-\alpha/4}$ with some constant c independent of n. Combining this with (6.6), we obtain a similar bound for the integral in (6.4), and taking into account (6.5), we get $J_1 \le cT^{-\alpha/8} + \varepsilon_n$. This completes the proof of Lemma 6.1. □

7. Binomial decomposition of convolutions

To show that the assumption $\limsup_{n\to\infty} I(Z_n) < \infty$ in Lemma 6.1 holds as long as $I(Z_{n_0}) < \infty$ for some n_0, we introduce a special decomposition of densities of Z_n. It is needed for the case $0 < \alpha < 2$, so this will be assumed below. Moreover, let $Z_n \Rightarrow Z$ weakly in distribution, where Z has a non-extremal stable law with index α.

To simplify the argument, assume $n_0 = 1$, so that $I(p) = I(X_1) < \infty$, where p denotes the density of X_1. In fact, we only consider the shifted normalized sums

$$\tilde{Z}_n = Z_n + a_n = \frac{X_1 + \cdots + X_n}{b_n},$$

and for the notational convenience, denote their densities by p_n. Note that, by the translation invariance, $I(Z_n) = I(\tilde{Z}_n)$.

Keeping the same notations as in the previous sections, we use a suitable truncation (which is actually not needed in case $\alpha > 1$). Introduce the probability densities

$$\tilde{p}_n(x) = \frac{b_n}{1-\delta_n} p(b_n x) 1_{\{|x| \le 1\}}, \qquad \tilde{q}_n(x) = \frac{b_n}{\delta_n} p(b_n x) 1_{\{|x|>1\}}$$

together with their characteristic functions

$$\tilde{f}_n(t) = \frac{1}{1-\delta_n} \int_{-b_n}^{b_n} \mathrm{e}^{\mathrm{i}tx/b_n} p(x)\,\mathrm{d}x, \qquad \tilde{g}_n(t) = \frac{1}{\delta_n} \int_{|x|>b_n} \mathrm{e}^{\mathrm{i}tx/b_n} p(x)\,\mathrm{d}x,$$

where $\delta_n = \int_{|x|>b_n} p(x)\,\mathrm{d}x$. Recall that $\delta_n \sim \frac{c}{n}$ with some constant $c > 0$, as emphasized in Proposition 5.3, cf. (5.9).

Then we have a binomial decomposition for convolutions

$$p_n = \left((1-\delta_n)\tilde{p}_n + \delta_n \tilde{q}_n\right)^{n*} = \sum_{k=0}^{n} \binom{n}{k} (1-\delta_n)^k \delta_n^{n-k} \tilde{p}_n^{k*} * \tilde{q}_n^{(n-k)*}. \tag{7.1}$$

Note that each convolution $\tilde{p}_n^{k*} * \tilde{q}_n^{(n-k)*}$ appearing in this weighted sum represents a probability density with characteristic function $\tilde{f}_n(t)^k \tilde{g}_n(t)^{n-k}$.

In this section, we establish some properties of $\tilde{f}_n$, which will be needed in the proof of Theorem 1.1. The corresponding density $\tilde{p}_n$ is supported on $[-1, 1]$, however, it does not need to have mean zero. So, put

$$d_n = \int_{-1}^{1} x \tilde{p}_n(x)\,\mathrm{d}x = \frac{1}{b_n(1-\delta_n)} \int_{-b_n}^{b_n} x p(x)\,\mathrm{d}x$$

and define

$$\psi_n(t) = \mathrm{e}^{-\mathrm{i}td_n} \tilde{f}_n(t),$$

which is the characteristic function of the centered random variable $\xi - d_n$, when ξ has density $\tilde{p}_n$. Thus, ψ_n corresponds to the density $r_n(x) = \tilde{p}_n(x + d_n)$, with $\psi_n'(0) = 0$.

The next two lemmas do not use the assumption $I(p) < \infty$ and may be stated for general distributions from the domain of attraction of these stable laws.

Lemma 7.1. *For all real t, with some constant C depending only on p,*

$$\left|\psi_n'(t)\right| \le \frac{C}{n}|t|. \tag{7.2}$$

Proof. The characteristic function ψ_n corresponds to the density $\tilde{p}_n(x+d_n)$. Using the property $\psi_n'(0)=0$, one may write

$$\begin{aligned}
\psi_n'(t) &= \int_{-1}^{1} \mathrm{i}(x-d_n)\big(\mathrm{e}^{\mathrm{i}t(x-d_n)}-1\big)\tilde{p}_n(x)\,\mathrm{d}x \\
&= \frac{\mathrm{i}b_n}{1-\delta_n}\int_{-1}^{1}(x-d_n)\big(\mathrm{e}^{\mathrm{i}t(x-d_n)}-1\big)p(b_n x)\,\mathrm{d}x \\
&= \frac{i}{1-\delta_n}\int_{-b_n}^{b_n}\left(\frac{x}{b_n}-d_n\right)\big(\mathrm{e}^{\mathrm{i}t(x/b_n-d_n)}-1\big)\,\mathrm{d}F_1(x),
\end{aligned}$$

where F_1 is the distribution function of X_1. Using $|\mathrm{e}^{is}-1|\le|s|$ ($s\in\mathbf{R}$), we deduce obvious upper bounds

$$\begin{aligned}
\big|\psi_n'(t)\big| &\le \frac{|t|}{1-\delta_n}\int_{-b_n}^{b_n}\left(\frac{x}{b_n}-d_n\right)^2\mathrm{d}F_1(x) \\
&\le \frac{2|t|}{b_n^2(1-\delta_n)}\int_{-b_n}^{b_n}x^2\,\mathrm{d}F_1(x)+\frac{2|t|}{1-\delta_n}d_n^2.
\end{aligned}$$

Integrating by parts, we have

$$\begin{aligned}
\int_{-b_n}^{b_n}x^2\,\mathrm{d}F_1(x) &= -b_n^2\big(1-F_1(b_n)+F_1(-b_n)\big)+2\int_0^{b_n}x\big(1-F_1(x)+F_1(-x)\big)\,\mathrm{d}x \\
&\le 2\int_0^{b_n}x\big(1-F_1(x)+F_1(-x)\big)\,\mathrm{d}x
\end{aligned}$$

and similarly

$$\begin{aligned}
\int_{-b_n}^{b_n}|x|\,\mathrm{d}F_1(x) &= -b_n\big(1-F_1(b_n)+F_1(-b_n)\big)+\int_0^{b_n}\big(1-F_1(x)+F_1(-x)\big)\,\mathrm{d}x \\
&\le \int_0^{b_n}\big(1-F_1(x)+F_1(-x)\big)\,\mathrm{d}x.
\end{aligned}$$

Since $1-\delta_n\to 1$, we get

$$\begin{aligned}
\big|\psi_n'(t)\big| &\le \frac{C|t|}{b_n^2}\int_0^{b_n}x\big(1-F_1(x)+F_1(-x)\big)\,\mathrm{d}x \\
&\quad+\frac{C|t|}{b_n^2}\left(\int_0^{b_n}\big(1-F_1(x)+F_1(-x)\big)\,\mathrm{d}x\right)^2
\end{aligned}\tag{7.3}$$

with some constant C depending on p.

Recall that in the asymptotical formulas (1.3)–(1.4) for F_1, the function B is equivalent to the slowly varying function h associated with the characteristic function of X_1. Thus, with some $c_0 \geq 0$, $c_1 \geq 0$ ($c_0 + c_1 > 0$), we have

$$F_1(x) = \frac{c_0 + \mathrm{o}(1)}{(-x)^\alpha} h(-x), \qquad x < 0; \qquad F_1(x) = 1 - \frac{c_1 + \mathrm{o}(1)}{x^\alpha} h(x), \qquad x > 0.$$

Hence, up to a constant, the first integral in (7.3) does not exceed

$$\int_0^{b_n} \frac{h(x)}{x^{\alpha-1}}\, \mathrm{d}x = b_n^{2-\alpha} h(b_n) \int_0^1 \frac{h(sb_n)}{h(b_n)} \frac{\mathrm{d}s}{s^{\alpha-1}}.$$

But, by the well-known result on slowly varying functions ([12], pages 66–67),

$$\int_0^1 \frac{h(sb_n)}{h(b_n)} \frac{\mathrm{d}s}{s^{\alpha-1}} \to \int_0^1 \frac{\mathrm{d}s}{s^{\alpha-1}} = \frac{1}{2-\alpha} \qquad \text{as } n \to \infty.$$

Therefore, with some constants C_1, C_2,

$$\frac{1}{b_n^2} \int_0^{b_n} x\big(1 - F_1(x) + F_1(-x)\big)\, \mathrm{d}x \leq C_1 b_n^{-\alpha} h(b_n) \leq \frac{C_2}{n},$$

where we have applied equation (5.8) of Proposition 5.3, telling us that $h(b_n) \sim b_n^\alpha / n$.

Now, consider the second integral in (7.3). In case $\alpha < 1$, again by [12], applied to the value $\alpha + 1$,

$$\int_0^1 \frac{h(sb_n)}{h(b_n)} \frac{\mathrm{d}s}{s^\alpha} \longrightarrow \int_0^1 \frac{\mathrm{d}s}{s^\alpha} = \frac{1}{1-\alpha} \qquad \text{as } n \to \infty.$$

Hence, using the asymptotic for F_1, the second integral in (7.3) does not exceed, up to a constant,

$$\int_0^{b_n} \frac{h(x)}{x^\alpha}\, \mathrm{d}x = b_n^{1-\alpha} h(b_n) \int_0^1 \frac{h(sb_n)}{h(b_n)} \frac{\mathrm{d}s}{s^\alpha} \sim \frac{b_n}{(1-\alpha)n}.$$

As a result,

$$\frac{1}{b_n^2} \left(\int_0^{b_n} \big(1 - F_1(x) + F_1(-x)\big)\, \mathrm{d}x \right)^2 \leq \frac{C_3}{n^2}$$

with some constant C_3, depending on p and α.

The case $1 < \alpha < 2$ is simpler, since then

$$\int_0^\infty \big(1 - F_1(x) + F_1(-x)\big)\, \mathrm{d}x < \infty,$$

while the factor $\frac{1}{b_n^2}$ behaves like $n^{-2/\alpha}$ (up to a slowly growing sequence), so it decays faster than $1/n$.

Finally, in case $\alpha = 1$, using the bound $h(x) \le C_\varepsilon x^\varepsilon$, $x \ge 1$ (where $\varepsilon > 0$ is any prescribed number), we see that, for large n the second integral in (7.3) does not exceed, up to a constant,

$$1 + \int_1^{b_n} \frac{h(x)}{x}\,\mathrm{d}x \le C b_n^{1/4}.$$

This yields

$$\frac{1}{b_n^2}\left(\int_0^{b_n} \bigl(1 - F_1(x) + F_1(-x)\bigr)\,\mathrm{d}x\right)^2 \le \frac{C}{b_n^{3/2}}$$

with some constant C depending on the density p. But the ratio $\frac{C}{b_n^{3/2}}$ behaves like $n^{-3/2}$ up to a slowly growing sequence, so it decays faster than $\frac{1}{n}$, as well. Thus, in all cases

$$\frac{1}{b_n^2}\left(\int_0^{b_n} \bigl(1 - F_1(x) + F_1(-x)\bigr)\,\mathrm{d}x\right)^2 = \mathrm{O}\left(\frac{1}{n}\right).$$

Lemma 7.1 is proved. □

Lemma 7.2. *Let $\delta \in (0, \alpha)$ and $\eta \in (0, 1)$ be fixed. There exist positive constants ε, c, C, depending on p, δ, η, with the following property: if $k \ge \eta n$, then*

$$\bigl|\psi_n(t)\bigr|^k = \bigl|\tilde{f}_n(t)\bigr|^k \le C\mathrm{e}^{-c|t|^\delta} \qquad \text{for } |t| \le \varepsilon b_n. \tag{7.4}$$

Proof. This is an analogue of the bound (5.10) for the characteristic functions of Z_n. In order to prove this upper bound, assume $|t| \ge 1$ and note that

$$\tilde{f}_n(t) = \frac{1}{1 - \delta_n}\bigl(f_1(t/b_n) - \delta_n \tilde{g}_n(t)\bigr), \qquad t \in \mathbf{R}. \tag{7.5}$$

To proceed, we apply Proposition 5.3. First recall that, according to Karamata's theorem, any positive slowly varying function $h(x)$ defined in $x \ge 0$ has a representation

$$h(x) = c(x)\exp\left\{\int_{x_0}^x \frac{w(y)}{y}\,\mathrm{d}y\right\}, \qquad x \ge x_0,$$

where $x_0 > 0$, $c(x) \to 1$, and $w(x) \to 0$, as $x \to \infty$. For $x_0 = \min_{n \ge 1} b_n$, $1 \le |t| \le \varepsilon b_n$, where $0 < \varepsilon \le 1$ is fixed, this representation implies that with some constant $c_0 > 0$

$$\frac{h(b_n/|t|)}{h(b_n)} \ge c_0 |t|^{-\gamma} \qquad \text{with } \gamma = \gamma(\varepsilon) = \sup_{y \ge 1/\varepsilon} \bigl|w(y)\bigr|.$$

Hence, from (5.7)–(5.8)

$$\bigl|f_1(t/b_n)\bigr| = \exp\bigl\{-c|t|^\alpha b_n^{-\alpha} h\bigl(b_n/|t|\bigr)\bigr\} \le \exp\bigl\{-c_1|t|^{\alpha-\gamma}/n\bigr\}$$

with some constant $c_1 > 0$.

We choose $\varepsilon > 0$ to be small enough so that $\gamma < \alpha - \delta$. Now, applying the above estimate in (7.5), we get in the region $1 \le |t| \le \varepsilon b_n$

$$\begin{aligned} |\tilde{f}_n(t)| &\le \frac{1}{1-\delta_n}\big(|f_1(t/b_n)| + \delta_n\big) \\ &\le \frac{1}{1-\delta_n}\big(\exp\{-c_1|t|^{\alpha-\gamma}/n\} + \delta_n\big). \end{aligned}$$

One can simplify the right-hand side by noting that $\frac{c_1|t|^{\alpha-\gamma}}{n} \le \frac{c_1 b_n^{\alpha-\gamma}}{n} < K$ with some constant K. Using $\log x \le x - 1$ ($x > 0$) and $\mathrm{e}^{-x} \le 1 - \frac{1}{K}(1 - \mathrm{e}^{-K})x$, for $0 \le x \le K$, we then have

$$\begin{aligned} \log\big(\exp\{-c_1|t|^{\alpha-\gamma}/n\} + \delta_n\big) &\le \exp\{-c_1|t|^{\alpha-\gamma}/n\} + \delta_n - 1 \\ &\le -\frac{1-\mathrm{e}^{-K}}{K}\frac{c_1|t|^{\alpha-\gamma}}{n} + \delta_n \\ &\le \frac{c_2}{n} - \frac{c_3|t|^{\alpha-\gamma}}{n} \end{aligned}$$

with positive constants c_j. As a result,

$$|\tilde{f}_n(t)| \le \exp\left\{\frac{1}{n}\big(c_4 - c_5|t|^{\alpha-\gamma}\big)\right\}$$

with some other positive constants c_4 and c_5 (independent of n). It remains to raise this inequality to the power k, and (7.4) follows. □

We will now develop a few applications of Lemmas 7.1 and 7.2 using the assumption $I(p) < \infty$. The latter forces p to have bounded variation and vanish at infinity. Hence,

$$\|r_n\|_{\mathrm{TV}} = \|\tilde{p}_n\|_{\mathrm{TV}} = b_n(1-\delta_n)^{-1}\|p1_{\{|x|\le b_n\}}\|_{\mathrm{TV}} \le b_n(1-\delta_n)^{-1}\sqrt{I(p)}. \tag{7.6}$$

Using the inequality (3.2), we see that the characteristic functions of $\tilde{p}_n$ and of the centered density $r_n(x) = \tilde{p}_n(x + d_n)$ satisfy

$$|\psi_n(t)| = |\tilde{f}_n(t)| \le \frac{cb_n}{|t|} \qquad (t \ne 0) \tag{7.7}$$

with some constant $c = c(p)$, depending on p, only.

Corollary 7.3. *If $I(p) < \infty$, then under the assumptions of Lemma 7.2 with $k \ge 4$, we have with some constant C depending on p, δ, η, only,*

$$\int_{-\infty}^{\infty}\big(1+|t|\big)\big|\psi_n^k(t)\big|\,\mathrm{d}t \le C, \tag{7.8}$$

$$\int_{-\infty}^{\infty} t^2\big|\big(\psi_n^k\big)'(t)\big|^2\,\mathrm{d}t \le C. \tag{7.9}$$

Proof. We have $(\psi_n^k)'(t) = k\psi_n'(t)\psi_n(t)^{k-1}$, while by (7.2),

$$\int_{-\infty}^{\infty} t^2 \left|\psi_n'(t)\right|^2 \left|\psi_n(t)\right|^{2(k-1)} \, \mathrm{d}t \le \frac{C^2}{n^2} \int_{-\infty}^{\infty} t^4 \left|\psi_n(t)\right|^{2(k-1)} \, \mathrm{d}t.$$

To estimate the last integral, first we use (7.4) which gives

$$\int_{|t| \le \varepsilon b_n} t^4 \left|\psi_n(t)\right|^{2(k-1)} \, \mathrm{d}t \le C.$$

For the complementary region $|t| > \varepsilon b_n$, note that

$$\tilde{f}_n(b_n t) = \frac{1}{1 - \delta_n} \int_{-b_n}^{b_n} \mathrm{e}^{\mathrm{i}tx} p(x) \, \mathrm{d}x,$$

which shows that these functions are separated from 1 uniformly in n in $|t| \ge \varepsilon$. (This can easily be seen by using general separation bounds for characteristic functions which are discussed in [2].) Thus,

$$\sup_{|t| \ge \varepsilon} \left|\psi_n(b_n t)\right| = \sup_{|t| \ge \varepsilon} \left|\tilde{f}_n(b_n t)\right| \le \mathrm{e}^{-c}$$

for some constant $c > 0$ independent of n. In addition, by (7.7),

$$t^4 \left|\psi_n(b_n t)\right|^6 \le \frac{c}{t^2}$$

with some other constant. Hence,

$$\int_{|t| \ge \varepsilon b_n} t^4 \left|\psi_n(t)\right|^{2(k-1)} \, \mathrm{d}t \le b_n^5 \mathrm{e}^{-2c(k-4)} \int_{|t| \ge \varepsilon} t^4 \left|\psi_n(b_n t)\right|^6 \, \mathrm{d}t \le C b_n^5 \mathrm{e}^{-2ck}.$$

The last expression is exponentially small with respect to n by the constraint on k, and we arrive at (7.9). The first inequality (7.8), which is simpler, is proved similarly. □

8. Boundedness of Fisher information. Proof of Theorem 1.1

In this section, we complete the last step in the proof of Theorem 1.1. Keeping the same notations as in the previous sections and recalling Lemma 6.1, we only need the following lemma.

Lemma 8.1. *Assume that $Z_n \Rightarrow Z$ weakly in distribution, where Z has a non-extremal stable law. If $I(Z_{n_0}) < \infty$ for some n_0, then $\sup_{n \ge n_0} I(Z_n) < \infty$.*

In the normal case, when X_1 has a finite second moment, the assertion immediately follows from Stam's inequality (2.3). In view of Lemma 6.1, we therefore obtain Barron–Johnson theorem, that is, $I(Z_n \| Z) \to 0$. Thus, we may focus on the case $0 < \alpha < 2$.

To simplify the argument and the notations, we assume $n_0 = 1$ (otherwise, mild modifications connected with the binomial decomposition are only needed). Thus, let $I(p) < \infty$, where p is the density of X_1. As in the previous section, we denote by p_n the density of $\tilde{Z}_n = Z_n + a_n$ and assume that $Z_n \Rightarrow Z$ weakly in distribution, where Z has a non-extremal stable law.

By Stam's inequality (2.3),

$$I(Z_n) \le \frac{b_n^2}{n} I(p).$$

Although the right-hand side tends to infinity, as $n \to \infty$, this inequality may be used for small values of n, and here it will be sufficient to show that $\sup_{n \ge n_0} I(Z_n) < \infty$ for some n_0.

Our basic tool is the binomial decomposition (7.1) of the previous section. Note that, by the convexity of the I-functional,

$$I(p_n) \le \sum_{k=0}^{n} \binom{n}{k} (1 - \delta_n)^k \delta_n^{n-k} I\bigl(\tilde{p}_n^{k*} * \tilde{q}_n^{(n-k)*}\bigr), \tag{8.1}$$

so it will be sufficient to properly estimate the terms in this sum. To this aim, we fix a number $\eta \in (0, 1)$ and distinguish two cases.

Lemma 8.2. *If $k \le n - 3$, then*

$$I\bigl(\tilde{p}_n^{k*} * \tilde{q}_n^{(n-k)*}\bigr) \le C(nb_n)^2 I(p) \tag{8.2}$$

with some constant C depending on p, only.

Proof. By the monotonicity property (2.2), $I(\tilde{p}_n^{k*} * \tilde{q}_n^{(n-k)*}) \le I(\tilde{q}_n^{(n-k)*})$. On the other hand, by Proposition 3.2, if $n - k \ge 3$,

$$I\bigl(\tilde{q}_n^{(n-k)*}\bigr) \le \tfrac{1}{2}\bigl(\bigl\|\tilde{q}_n^{[(n-k)/3]*}\bigr\|_{\mathrm{TV}}^2 + 2\bigl\|\tilde{q}_n^{[(n-k)/3]*}\bigr\|_{\mathrm{TV}} \cdot \bigl\|\tilde{q}_n^{n-k-2[(n-k)/3]*}\bigr\|_{\mathrm{TV}}\bigr).$$

But the total variation norm decreases when taking convolutions, so that $\|\tilde{q}_n^{s*}\|_{\mathrm{TV}} \le \|\tilde{q}_n\|_{\mathrm{TV}}$ $(s = 1, 2, \dots)$. Hence,

$$I\bigl(\tilde{q}_n^{(n-k)*}\bigr) \le \tfrac{3}{2}\|\tilde{q}_n\|_{\mathrm{TV}}^2.$$

In turn, by means of the inequality $\|p\|_{\mathrm{TV}} \le \sqrt{I(p)}$ (Proposition 3.1), we have

$$\|\tilde{q}_n\|_{\mathrm{TV}} = b_n \delta_n^{-1} \|p 1_{\{|x| > b_n\}}\|_{\mathrm{TV}} \le b_n \delta_n^{-1} \|p\|_{\mathrm{TV}} \le b_n \delta_n^{-1} \sqrt{I(p)},$$

where we used the property $p(-\infty) = p(\infty) = 0$ for the first inequality. Thus

$$I\bigl(\tilde{p}_n^{k*} * \tilde{q}_n^{(n-k)*}\bigr) \le \tfrac{3}{2}\bigl(\sqrt{I(p)} b_n \delta_n^{-1}\bigr)^2.$$

Recalling that $\delta_n \sim \frac{c}{n}$, Lemma 8.2 is proved. □

Lemma 8.3. *If* $15 \le \eta n \le k \le n$, *then*

$$I\big(\tilde{p}_n^{k*} * \tilde{q}_n^{(n-k)*}\big) \le C \tag{8.3}$$

with some constant C *depending on* p *and* η, *only.*

Proof. Again appealing to the monotonicity of the Fisher information, we will use the bound

$$I\big(\tilde{p}_n^{k*} * \tilde{q}_n^{(n-k)*}\big) \le I\big(\tilde{p}_n^{k*}\big).$$

Thus, involving the centered density $r_n(x) = \tilde{p}_n(x + d_n)$ with the characteristic function ψ_n (as in the previous section), it suffices to show that

$$I\big(r_n^{k*}\big) = I\big(\tilde{p}_n^{k*}\big) \le C. \tag{8.4}$$

Assume first that $\eta_0 n \le k \le n$, where $0 < \eta_0 < \eta$. Since $\|r_n\|_{\mathrm{TV}} \le C b_n \sqrt{I(p)} < \infty$ (see (7.6) and Proposition 3.2), the convolution powers r_n^{k*} have finite Fisher information, whenever $k \ge 3$. In view of the bound (7.7) on the characteristic functions, we may invoke inversion formulas like in (5.5)–(5.7) to write, for any $x \in \mathbf{R}$,

$$r_n^{k*}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}tx} \psi_n(t)^k \,\mathrm{d}t, \tag{8.5}$$

$$\big(r_n^{k*}\big)'(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}tx} (-\mathrm{i}t) \psi_n(t)^k \,\mathrm{d}t, \tag{8.6}$$

$$r_n^{k*}(x) + x\big(r_n^{k*}\big)'(x) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}tx} tk \psi_n(t)^{k-1} \psi_n'(t) \,\mathrm{d}t, \tag{8.7}$$

where for reasons of integrability it is safer to assume that $k \ge 5$.

Corollary 7.3 tells us that the Fourier transforms in (8.5) and (8.7) are well defined for square integrable functions whose L^2-norms are bounded by a constant independent of k and n. Hence, the same is true for

$$x\big(r_n^{k*}\big)'(x) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}tx} \big(\psi_n(t)^k + tk \psi_n(t)^{k-1} \psi_n'(t)\big) \,\mathrm{d}t,$$

and we may write

$$\big|\big(r_n^{k*}\big)'(x)\big| \le \frac{u_{nk}(x)}{|x|} \tag{8.8}$$

with

$$\|u_{nk}\|_2^2 = \int_{-\infty}^{\infty} u_{nk}(x)^2 \,\mathrm{d}t \le C. \tag{8.9}$$

Moreover, according to (7.8), L^1-norms of the functions $(-\mathrm{i}t)\psi_n(t)^k$ in (8.6) are also bounded by a constant independent of k and n. Hence,

$$\sup_x \big|\big(r_n^{k*}\big)'(x)\big| \le C$$

for all n and $\eta_0 n \le k \le n$. As a result, (8.8) may be sharpened to

$$\big|\big(r_n^{k*}\big)'(x)\big| \le \frac{u_{nk}(x)}{1+|x|}$$

with some functions u_{nk} satisfying (8.9). By applying Cauchy's inequality, the latter immediately implies that

$$\|r_n^{k*}\|_{\mathrm{TV}} = \int_{-\infty}^{\infty} \big|\big(r_n^{k*}\big)'(x)\big|\,\mathrm{d}x \le C'\|u_{nk}\|_2 \le C, \tag{8.10}$$

where the resulting constant C may depend on p and η_0 (by choosing, e.g., $\delta = \alpha/2$ in the previous auxiliary lemmas of the previous section).

We now apply Proposition 3.2 to convolutions of any three densities r_n^{k*}, as above. That is, if $\eta_0 n \le k_j \le n$ and $k_j \ge 5$ ($j = 1, 2, 3$), we obtain by (3.4) and (8.10) that

$$I\big(r_n^{(k_1+k_2+k_3)*}\big) \le \tfrac{3}{2}C^2. \tag{8.11}$$

Starting with $k \ge 15$, put $k_1 = k_2 = [\frac{k}{3}]$, $k_3 = k - (k_1 + k_2)$, so that $k_j \ge 5$. Also, if $k \ge \eta n$, we have $k_j \ge [\frac{\eta n}{3}] \ge \frac{\eta n}{6}$. Hence, we may choose $\eta_0 = \frac{\eta}{6}$, and thus (8.11) implies (8.3)–(8.4). $\square$

Proof of Lemma 8.1. In the case $15 \le \eta n \le n - 3$, we may combine Lemmas 8.2 and 8.3 to get from (8.1) the following. With some constant $C = C(p, \eta)$, depending on η and the density p via $I(p)$ and the constant c in $\delta_n \sim \frac{c}{n}$,

$$\begin{aligned} I(p_n) &\le C(nb_n)^2 \sum_{0\le k<\eta n} \binom{n}{k} (1-\delta_n)^k \delta_n^{n-k} + C \sum_{\eta n \le k \le n} \binom{n}{k} (1-\delta_n)^k \delta_n^{n-k} \\ &\le C(nb_n)^2 \cdot 2^n \delta_n^{(1-\eta)n} + C \le C', \end{aligned}$$

where the last inequality holds for all sufficiently large n (by using $\delta_n \sim \frac{c}{n}$) with, for example, $\eta = \frac{1}{2}$. Lemma 8.1 and therefore Theorem 1.1 are now proved. $\square$

Remark 8.4. Finally, let us comment on the conditions (a)–(b) from the Introduction. In view of the general bound (3.3), (a) is always necessary for the finiteness of $I(Z_n)$ with some n. Since (b) is weaker than (a), we need explain the opposite direction.

If $1 < \alpha \le 2$, then X_1 has finite first absolute moment $C = \mathbf{E}|X_1|$. Hence, under (1.6), the condition (3.5) is fulfilled and thus the bound (3.6) is applicable to all Z_n with $n \ge (\nu + 2)/2$. More precisely, denoting by $g_n(t) = f_1(t)^n$ the characteristic function of $S_n = X_1 + \cdots + X_n$, we have

$$\big|\big(tg_n(t)\big)'\big| \le \big|g_n(t)\big| + |t|\big|g_n'(t)\big| \le \big(1 + Cn|t|\big)\big|f_1(t)\big|^{n-1},$$

thus S_n has a density $\rho_n(x)$ whose total variation norm satisfies

$$\|\rho_n\|_{\mathrm{TV}}^4 \le \int_{-\infty}^{\infty} t^2\big|f_1(t)\big|^{2n}\,\mathrm{d}t \int_{-\infty}^{\infty} \big(1 + Cn|t|\big)^2\big|f_1(t)\big|^{2(n-1)}\,\mathrm{d}t < \infty.$$

By Proposition 3.2, we get $I(S_{3n}) < \infty$.

Acknowledgements

Research partially supported by NSF Grant DMS-11-06530, Simons Foundation and SFB 701. We thank the referees for careful reading of the manuscript and valuable comments.

References

[1] Barron, A.R. (1986). Entropy and the central limit theorem. *Ann. Probab.* **14** 336–342. MR0815975
[2] Bobkov, S.G., Chistyakov, G.P. and Götze, F. (2012). Bounds for characteristic functions in terms of quantiles and entropy. *Electron. Commun. Probab.* **17** 1–9. MR2943104
[3] Bobkov, S.G., Chistyakov, G.P. and Götze, F. (2013). Fisher information and the central limit theorem. *Probab. Theory Related Fields*. Published online: 23 April 2013.
[4] Bobkov, S.G., Chistyakov, G.P. and Götze, F. (2013). Convergence to Stable Laws in Relative Entropy. *J. Theoret. Probab.* **26** 803–818. MR3090551
[5] Csiszár, I. (1967). Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.* **2** 299–318. MR0219345
[6] Ibragimov, I.A. and Linnik, J.V. (1965). *Nezavisimye Stalionarno Svyazannye Velichiny*. Moscow: Nauka. MR0202176
[7] Johnson, O. (2004). *Information Theory and the Central Limit Theorem*. London: Imperial College Press. MR2109042
[8] Johnson, O. and Barron, A. (2004). Fisher information inequalities and the central limit theorem. *Probab. Theory Related Fields* **129** 391–409. MR2128239
[9] Kullback, S. (1967). A lower bound for discrimination in terms of variation. *IEEE Trans. Inform. Theory* **T-13** 126–127.
[10] Linnik, J.V. (1959). An information-theoretic proof of the central limit theorem with Lindeberg conditions. *Theory Probab. Appl.* **4** 288–299. MR0124081
[11] Pinsker, M.S. (1964). *Information and Information Stability of Random Variables and Processes. Translated and Edited by Amiel Feinstein*. San Francisco, CA: Holden-Day. MR0213190
[12] Seneta, E. (1976). *Regularly Varying Functions*. *Lecture Notes in Math.* **508**. Berlin: Springer. MR0453936
[13] Stam, A.J. (1959). Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Information and Control* **2** 101–112. MR0109101
[14] Yamazato, M. (1978). Unimodality of infinitely divisible distribution functions of class L. *Ann. Probab.* **6** 523–531. MR0482941
[15] Zolotarev, V.M. (1983). *One-dimensional stable distributions*. *Probability Theory and Mathematical Statistics*. Moscow: Nauka.

Received August 2012 and revised April 2013